Fréchet Regression on the Bures-Wasserstein Manifold

Anonymous Authors

Abstract

Fréchet regression, or conditional Barycenters, is a flexible framework for modeling relationships between covariates (usually Euclidean) and response variables on general metric spaces, e.g., probability distributions or positive definite matrices. However, in contrast to classical barycenter problems, computing conditional counterparts in many non-Euclidean spaces remains an open challenge, as they yield non-convex optimization problems with an affine structure. In this work, we study the existence and computation of conditional barycenters, specifically in the manifold $\mathbb{S}_{++}^d$ of positive-definite matrices with the Bures-Wasserstein metric.

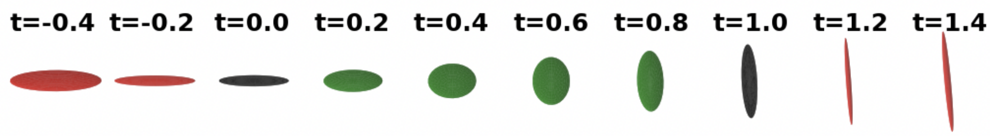

Figure 1: Interpolation (green) and extrapolation (red) between two (black) ellipsoids (3x3 SPD matrices) under BW metric.

Let $\Sigma_k \in \mathbb{S}_{++}^d$ with $\sum_{k=1}^n \lambda_k = 1$, we aim to find

$$\min_{S \in \mathbb{S}_{++}^d} F(S) := \sum_{k=1}^n \lambda_k W_2^2(S, \Sigma_k) = \sum_{i \in \mathcal{I}} \lambda_i^+ W_2^2(S, \Sigma_i) - \sum_{j \in \mathcal{J}} \lambda_j^- W_2^2(S, \Sigma_j), \tag{1}$$

for $\lambda_i^+, \lambda_j^- > 0$, $\mathcal{I} = \{k : \lambda_k > 0\}$, $\mathcal{J} = \{k : \lambda_k < 0\}$. Without specific constraints, negative weights can drive the solution to the boundary of the cone or to infinity. We provide a sufficient condition for the existence of a minimizer of the conditional barycenter problem that characterizes the regression range of extrapolation.

Theorem 1 (**Spectral Dominance of Positive Weights**). *Let $\Sigma_k \in \mathbb{S}_{++}^d$ with $\sum_{k=1}^n \lambda_k = 1$. If*

$$\sum_{i \in \mathcal{I}} \lambda_i^+ \sqrt{\lambda_{\min}(\Sigma_i)} > \sum_{j \in \mathcal{J}} \lambda_j^- \sqrt{\lambda_{\max}(\Sigma_j)}, \tag{2}$$

then Problem (1) *admits a solution in $\mathbb{S}_{++}^d$. Furthermore, any stationary point S_* satisfies*

$$\Big(\sum_{i \in \mathcal{I}} \lambda_i^+ \sqrt{\lambda_{\min}(\Sigma_i)} - \sum_{j \in \mathcal{J}} \lambda_j^- \sqrt{\lambda_{\max}(\Sigma_j)}\Big)^2 I \prec S_* \prec \Big(\sum_{i \in \mathcal{I}} \lambda_i^+ \sqrt{\lambda_{\max}(\Sigma_i)} - \sum_{j \in \mathcal{J}} \lambda_j^- \sqrt{\lambda_{\min}(\Sigma_j)}\Big)^2 I.$$

Moreover, we further characterize the optimization landscape, proving that under this condition, the objective is free of local maxima. Additionally, we develop a projection-free and provably correct algorithm for the approximate computation of first-order stationary points. Finally, we provide a stochastic reformulation that enables the use of off-the-shelf stochastic Riemannian optimization methods for large-scale setups. Numerical experiments validate the performance of the proposed methods on regression problems of real-world biological networks and on large-scale synthetic Diffusion Tensor Imaging problems.